



Nathaniel Dean

Texas State University-San Marcos

Data, Unsolved Problems, Computing and Incomprehensibility

nd17@txstate.edu

To understand interactions between entities (for example, people, objects or groups) systems of interactions are often modeled as *graphs* linking *nodes* (entities) with *edges* (representing various types of connections between the entities). After data collection there are many statistical approaches to analyzing the data. Our approach is to model data as a graph and explore the graph using a variety of tools such as optimization and visualization. In this talk we discuss ways to construct graphs from data, and we show how to use the graphs to reveal patterns. For example, we propose a metric that ranks groups based on their interactions with others. It applies to a variety of applications, its mean and variance can be computed quickly, and hence extremists (for example) can be identified in real time and the results used in interactive data analysis environments. Applications are numerous including social network analysis, counterterrorism, market basket analysis, epidemiology and molecular biology. The limitations of this approach are also discussed explaining why some graphs cannot be visualized and hence why certain data cannot be understood.